Mon. Not. R. Astron. Soc. **000**, 1–13 (1998) Printed 21 May 2009 (MN LATEX style file v2.2)

Analysis of non-Gaussian CMB maps based on the N-pdf. Application to WMAP data

P. Vielva[1], J. L. Sanz[1,2]
[1] *Instituto de Física de Cantabria (CSIC - Univ. de Cantabria), Avda. Los Castros s/n, 39005 - Santander, Spain*
[2] *CNR Istituto de Scienza e Tecnologie dell'Informazione, 56124, Pisa, Italy*
E-mails : vielva@ifca.unican.es, sanz@ifca.unican.es

Accepted ???. Received ???; in original form 21 May 2009

ABSTRACT
We present a new method based on the N-point probability distribution (pdf) to study non-Gaussianity in cosmic microwave background (CMB) maps. Likelihood and Bayesian estimation are applied to a local non-linear perturbed model up to third order, characterized by a linear term which is described by a Gaussian N-pdf, and a second and third order terms which are proportional to the square and the cube of the linear one. We also explore a set of model selection techniques (the Akaike and the Bayesian Information Criteria, the minimum description length, the Bayesian Evidence and the Generalized Likelihood Ratio Test) and their application to decide whether a given data set is better described by the proposed local non-Gaussian model, rather than by the standard Gaussian temperature distribution. As an application, we consider the analysis of the WMAP 5-year data at a resolution of $\approx 2^\circ$. At this angular scale (the Sachs-Wolfe regime), the non-Gaussian description proposed in this work defaults (under certain conditions) to an approximative local form of the weak non-linear coupling inflationary model (e.g. Komatsu et al. 2001) previously addressed in the literature. For this particular case, we obtain an estimation for the non-linear coupling parameter of $-94 < f_{NL} < 154$ at 95% CL. Equally, model selection criteria also indicate that the Gaussian hypothesis is favored against the particular local non-Gaussian model proposed in this work. This result is in agreement with previous findings obtained for equivalent non-Gaussian models and with different non-Gaussian estimators. However, our estimator based on the N-pdf is more efficient than previous estimators and, therefore, provides tighter constraints on the coupling parameter at degree angular resolution.

Key words: cosmology: observations – cosmology: cosmic microwave background – methods: data analysis – methods: statistical

1 INTRODUCTION

The Cosmic Microwave Background (CMB) fluctuations, a relic radiation originated around 400,000 years after the Big-Bang, is one of the most outstanding sources for understanding the evolution and energy/matter content of the Universe. The large amount of high quality data provided by recent CMB experiments and other complementary astronomical observations, have provided with a consistent picture for a flat Universe filled with cold dark matter (CDM) and dark energy in the form of a cosmological constant (Λ), plus the standard baryonic and electromagnetic components: the *concordance model* (e.g. Komatsu et al. 2008). However, beyond the strength of the CMB measurements to put constraints on the cosmological parameters, like the ones already provided by the NASA *Wilkinson Microwave Anisotropy Probe* (WMAP, Hinshaw et al. 2008) and the ones expected from the incoming ESA *Planck satellite*, the CMB is a unique tool to probe fundamental principles and assumptions of the so-called *standard model*. In particular, the application of sophisticated statistical analysis to CMB data might help us to understand whether the temperature fluctuations of the primordial radiation are compatible with the fundamental isotropic and Gaussian predictions from the *inflationary phase*. The basic inflationary scenario relates the CMB fluctuations, as well as the large-scale structure of the Universe, to the Gaussian quantum energy density perturbations present during the early Universe (see for instance Liddle & Lyth 2000). The present homogeneity and isotropy of the Universe is compatible with an inflationary era in the early universe and this idea is the only way,

nowadays, to explain efficiently current observations ranging from galaxies to the CMB. In particular, this fundamental hypothesis predicts that the CMB fluctuations follow an isotropic and Gaussian random. In fact, the estimation of the cosmological parameters defining the concordance model is done by assuming these statistical properties.

The quality of current CMB data (in particular the ones provided by the WMAP satellite) has allowed for a systematic probe of the statistical properties of the relic radiation. Indeed, the interest of the scientific community in this field has experimented a significant growth, since the analysis of the WMAP data has reported several hints for departure from isotropy and Gaussianity of the CMB temperature distribution. Some of these works are the following. Park (2004) detected a Gaussianity deviation with a genus-based statistic; Vielva et al. (2004) found a significant non-Gaussian signature on the 1-year WMAP data on the kurtosis of the Spherical Mexican Hat (SMHW) wavelet coefficients at scales of around 10 degrees, pointing out a very large *cold spot* (CS) on the southern hemisphere as a possible source for this non-Gaussianity. A posterior analysis (Cruz et al. 2005) confirmed the anomalous nature of the CS by performing an area-based statistical analysis of the wavelet coefficients. This detection, as well as new ones, were confirmed by analyzing WMAP with various wavelet bases and several statistical estimators (Mukherjee & Wang 2004; Cayón et al. 2005; McEwen et al. 2005; Cruz et al. 2006; Martínez-González et al. 2006; Cruz et al. 2007a; McEwen et al. 2006; Pietrobon et al. 2008; Wiaux et al. 2008). Isotropy deviations were also reported in different manners: an anomalous alignment of the low multipoles of the CMB (Copi et al. 2004; de Oliveira-Costa et al. 2004; Katz & Weeks 2004; Schwarz et al. 2004; Bielewicz et al. 2005; Land & Magueijo 2005b; Abramo et al. 2006; Freeman et al. 2006; Land & Magueijo 2007); north-south asymmetries of the CMB fluctuations (Eriksen et al. 2004a,b; Hansen et al. 2004a,b; Donoghue & Donoghue 2005; Eriksen et al. 2005; Land & Magueijo 2005a; Bernui et al. 2006, 2007; Eriksen et al. 2007; Räth et al. 2007; Gordon 2007); an anomalous variance value (Monteserín et al. 2008); unexpected correlation among the CMB phases (Chiang et al. 2003; Coles et al. 2004; Chiang & Naselsky 2006); unbalanced distribution of the temperature extrema (Tojeiro et al. 2006); and anomalous alignment of CMB structures (Wiaux et al. 2006; Vielva et al. 2006, 2007).

All the previous analyses can be considered as *blind approaches*, since null tests were performed to probe the CMB compatibility with the isotropic and Gaussian hypotheses. Complementary to these works, the reader can also find in the literature several studies where *targeted* departures from Gaussianity are explored, based on non-standard physical models. In particular, several analyses have studied the WMAP data compatibility with anisotropic universes: the Bianchi VII_h model (Jaffe et al. 2006a,b,c; Bridges et al. 2007a, 2008). Recently, some works have proposed non-rotational invariant models like Böhmer & Mota (2008) and Ackerman et al. (2007) (exprored by Groeneboom & Eriksen 2008, in the context of WMAP data). Although these non-rotational invariant models are promising and provide us with anisotropic templates that could help to fix some anomalies in WMAP data, more work is still needed to connect these anisotropic patterns of the CMB fluctuations to a satisfactory physical model describing the evolution of the anisotropic field (Himmetoglu et al. 2008a,b).

In addition to the previous analyses, several works have studied different hypotheses to explore the anomalous nature of the CS; for instance Cruz et al. (2008) explored the CS compatibility with different non-standard models, pointing out that an explanation in terms of a cosmic texture (as proposed by Cruz et al. 2007b) is much more favored than other alternatives already discussed in the literature, like a very large void in the large scale structure of the Universe or contamination, in the form of the Sunyaev-Zeldovich emission, due to a large and nearby galaxy cluster. However, the study of non-standard inflationary models is the problem that has attracted a larger attention. For instance, and for the WMAP case, the non-linear coupling parameter $\mathrm{f_{NL}}$ that describes the non-linear evolution of the inflationary potential (see e.g. Bartolo et al. 2004, and references therein) has been constrained by several groups: using the angular bispectrum (Komatsu et al. 2003; Creminelli et al. 2006; Spergel et al. 2007; Komatsu et al. 2008); applying the Minkowski functionals (Komatsu et al. 2003; Spergel et al. 2007; Gott et al. 2007; Hikage et al. 2008; Komatsu et al. 2008); and using different statistics based on wavelets (Mukherjee & Wang 2004; Cabella et al. 2005; Curto et al. 2008). Besides a claim for $\mathrm{f_{NL}} > 0$ with a probability greater than 95% (Yadav & Wandelt 2008), there is a general consensus on the WMAP compatibility with the predictions made by the standard inflationary scenario at least at 95% confidence level. The current best limits (Curto et al. 2008) are: $-8 < \mathrm{f_{NL}} < 111$ at 95% CL.

The present paper is related to *non-blind* or *targeted* probes for Gaussianity deviations, and, more specifically, it addresses two common topics arose in these kind of studies: the definition of an optimal estimator for the non-standard model (and, therefore, the optimal estimation of the parameters that define such a model) and the complementary issue of model selection. The former aspect is, usually, quite complex, since, for many physical models and realistic observational limitations (e.g. incomplete sky coverage) there is not a trivial solution and, therefore, non-optimal estimators are usually proposed (which are posteriorly characterized by simulations). The latter issue on model selection is equally complex, since, generally, relies on heuristic principles. Some authors adopt a fully Bayesian view, whereas others adopt asymptotic measurements for the distance between two hypothesis, and others prefer to rely on the information that can be obtained from the likelihood itself. In this work we propose to study a non-Gaussian model for the CMB that is a local non-linear expansion of the temperature fluctuations. For this model —that, at large scales, is considered by some authors as an approximation to the non-linear coupling parameter $\mathrm{f_{NL}}$— we are able to build the exact likelihood on pixel space. To work in pixel space allows one to include easily non-ideal observational conditions, like incomplete sky coverage and anisotropic noise. We show how, from this likelihood, it is straightforward to obtain an analytical expression for the optimal estimator for the non-Gaussian term. In addition, we also explore different model selection criteria, like the Akaike information criterion (Akaike 1973), the Bayesian information criterion (Schwarz 1978), the min-

imum description length (Rissanen 2001), the Bayesian evidence, and the generalized likelihood ratio test.

The paper is organized as follows. In Section 2 we describe the physical model based on the local expansion of the CMB fluctuations and derive the full posterior probability. In Section 3 we address the issue of the parameter estimation (3.1) and model selection (3.2). The methodology is explored on WMAP-like simulations in Section 4 and it is applied to WMAP 5-year data in Section 5. Finally, conclusions are given in Section 6.

2 THE NON-GAUSSIAN MODEL

Current observations indicate that CMB temperature fluctuations can be well described by random Gaussian fluctuations, as predicted by the standard inflationary scenario. However, as it was discussed in the Introduction, these observations still allow for a small departure from the Gaussian distribution, which could reflect the role played by physical processes described by non-standard models for the structure formation.

In this paper, we will focus on a parametric non-Gaussian model, that accounts for a small and local (i.e. point-to-point) perturbation of the CMB temperature fluctuations, around its intrinsic Gaussian distribution:

$$\Delta T_i = (\Delta T_i)_G + a\left[(\Delta T_i)_G^2 - \left\langle (\Delta T_i)_G^2 \right\rangle\right] + b\,(\Delta T_i)_G^3\,. \quad (1)$$

$(\Delta T_i)_G$ (*the linear term*) is the Gaussian part, whose N-point probability density function (N-pdf) can be easily described in terms of the standard inflationary model. The second and third terms on the right-hand side are *the quadratic* and *cubic perturbation terms*, respectively. Their corresponding contribution to the observed CMB fluctuations (ΔT_i) is governed by the a and b parameters. Subindex i refers to the direction corresponding to a certain pixelization on the sphere, and the operator $\langle\cdot\rangle$ averages over all the pixels defining the sky coverage. Notice that the previous expression does not include instrumental noise. We have adopted this simplification, since we aim to focus the work on large-scale CMB data sets ($> 1^\circ$), where (as it is the case for WMAP), the contribution from noise is typically negligible.

Let us point out that the particular situation for $b \equiv 0$ defaults into a well known case that has been already addressed in the literature (e.g. Komatsu et al. 2001; Cayón et al. 2003; Curto et al. 2007). It describes a local approximation to the weak non-linear coupling inflationary model (e..g Komatsu et al. 2001; Liguori et al. 2003) at scales larger than the horizon scale at the recombination time (i.e. above the degree scale). In this context, the a factor is usally related to the non-linear coupling parameter $\mathrm{f_{NL}}$ by:

$$a \equiv \frac{3\mathrm{f_{NL}}}{T_0}, \quad (2)$$

where $T_0 = 2,725$ mK is the CMB temperature, and we follow the sign convention in Liddle & Lyth (2000) for the relation between the temperature fluctuations and the gravitational potential at the Sachs-Wolfe regime. Let us remark that, of course, this model do not pretend to incorporate all the gravitational (like lensing) and non-gravitational effects, due to the evolution of the initial quadratic potential model. Indeed, the reason to select the specific model given by equation 1 is twofold. One the one hand, it is an useful parametrization for describing a small departure from Gaussianity, that allows us to present a new methodology. On the other hand, it is a model previously addressed by other authors (using other estimators), and, therefore, it is easier to make a straightforward comparison among the results.

For simplicity, let us transform the Gaussian part $(\Delta T_i)_G$ into a zero mean and unity variance random variable ϕ_i, hence, equation 1 can be rewritten as:

$$x_i = \phi_i + \epsilon\left(\phi_i^2 - 1\right) + \alpha\epsilon^2\phi_i^3, \quad (3)$$

where:

$$x \equiv \frac{1}{\sigma}\Delta T, \quad \phi \equiv \frac{1}{\sigma}\left(\Delta T\right)_G, \quad \epsilon \equiv a\sigma, \quad \alpha \equiv \frac{b\sigma^2}{\epsilon^2}, \quad (4)$$

and $\sigma^2 \equiv \left\langle (\Delta T_i)_G^2 \right\rangle$ is the rms of the CMB fluctuations. Let us remark that, since the proposed non-Gaussian model is a perturbation of the standard Gaussian one, the non-linear parameters have to satisfy:

$$|\epsilon| \ll 1, \quad |\alpha| \lesssim 1. \quad (5)$$

It is straightforward to show that the normalized Gaussian field ϕ satisfies:

$$\langle\phi_i\rangle = \langle\phi_i^3\rangle = 0, \quad \langle\phi_i^2\rangle = 1, \quad \langle\phi_i\phi_j\rangle = \xi_{ij}, \quad (6)$$

where ξ_{ij} represents the normalized correlation between pixels i and j. The N-pdf of the $\boldsymbol{\phi} = \{\phi_1, \phi_2, ..., \phi_N\}$ random field (where N refers to the number of pixels on the sphere that are observed) is given by a multivariate Gaussian:

$$p(\boldsymbol{\phi}) = \frac{1}{(2\pi)^{N/2}(\det\boldsymbol{\xi})^{1/2}} e^{-\frac{1}{2}\boldsymbol{\phi}\boldsymbol{\xi}^{-1}\boldsymbol{\phi}^t}, \quad (7)$$

where $\boldsymbol{\xi}$ denotes the correlation matrix and operator $\cdot^t$ denotes standard matrix/vector transpose.

Our goal is to compute the N-pdf associated to the non-Gaussian $\boldsymbol{x} = \{x_1, x_2, ..., x_N\}$ field, as a function of the non-linear coupling parameters based on the full N-pdf for the underlaying Gaussian signal $\boldsymbol{\phi}$. Hence, let us make first the inversion of equation 3 to find the expression of ϕ_i as a function of x_i:

$$\phi_i = x_i - \epsilon(x_i^2 - 1) + \epsilon^2[(2-\alpha)x_i^3 - 2x_i] + O(3). \quad (8)$$

Obviously, since previous equation is a local transformation, the Jacobian matrix is diagonal and, therefore, the Jacobian (Z) is given by:

$$Z = \det\left[\frac{\partial\phi_i}{\partial x_j}\right] = \prod_i\left(\frac{\partial\phi_i}{\partial x_i}\right). \quad (9)$$

It is more convenient to work with the log-Jacobian ($\log Z$), which, taking into account equation 8 and expanding the log function up to second order, is given by:

$$\begin{aligned}\log Z &= \sum_i \log\left(\frac{\partial\phi_i}{\partial x_i}\right) \quad (10)\\ &= -2\epsilon\sum_i x_i + \epsilon^2\left[-2N + (4-3\alpha)\sum_i x_i^2\right] + 0(3).\end{aligned}$$

Now, taking into account equation 3 and recalling that $\boldsymbol{\phi}$ is a Gaussian field, it is straightforward to prove that the data

$\boldsymbol{x}$ satisfy: $\sum_i x_i = 0$ and $\frac{1}{N}\sum_i x_i^2 = 1 + O(2)$. Therefore, the log-Jacobian is given by:

$$\log Z = N\epsilon^2 (2 - 3\alpha) + O(3), \tag{11}$$

Finally, it is easy to calculate $p(\boldsymbol{x}|\epsilon)$, the N-point pdf of $\boldsymbol{x}$ given the ϵ parameter:

$$p(\boldsymbol{x}|\epsilon) = p(\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{x}))Z = p(\boldsymbol{x}|0)e^{l(\boldsymbol{x}|\epsilon)} \tag{12}$$

where the probability $p(\boldsymbol{x}|0)$ is given by:

$$p(\boldsymbol{x}|0) = \frac{1}{(2\pi)^{N/2}(\det(\boldsymbol{\xi}))^{1/2}} e^{-\frac{1}{2}\boldsymbol{x}\boldsymbol{\xi}^{-1}\boldsymbol{x}^t}, \tag{13}$$

i.e., it is the N-point Gaussian pdf (i.e. $p(\boldsymbol{x}|0) \equiv p(\boldsymbol{\phi} \equiv \boldsymbol{x})$), and the log-likelihood $l(\boldsymbol{x}|\epsilon)$ reads as:

$$l(\boldsymbol{x}|\epsilon) = \log Z - \frac{1}{2}\left[\phi\boldsymbol{\xi}^{-1}\phi^t - x\boldsymbol{\xi}^{-1}x^t\right]. \tag{14}$$

It is easy to show that $l(\boldsymbol{x}|\epsilon)$ is given by:

$$l(\boldsymbol{x}|\epsilon) \equiv N[\epsilon R - \epsilon^2 Q + O(3)], \tag{15}$$

where the functions R and Q are given by:

$$R = \frac{1}{N}\boldsymbol{x}\boldsymbol{\xi}^{-1}\left[\boldsymbol{x}^2 - \mathbb{I}\right]^t, \tag{16}$$

and

$$Q = -2 + J + \alpha S, \tag{17}$$

being $\mathbb{I}$ the unity vector of dimension N and

$$J = \frac{1}{N}\left\{2\boldsymbol{x}\boldsymbol{\xi}^{-1}\left[\boldsymbol{x}\left(\boldsymbol{x}^2 - \mathbb{I}\right)\right]^t + \frac{1}{2}\left[\boldsymbol{x}^2 - \mathbb{I}\right]\boldsymbol{\xi}^{-1}\left[\boldsymbol{x}^2 - \mathbb{I}\right]^t\right\} \tag{18}$$

and

$$S = \frac{3}{N}\left\{\mathbb{I}\boldsymbol{\xi}^{-1}\mathbb{I}^t - \frac{1}{3}\boldsymbol{x}\boldsymbol{\xi}^{-1}\boldsymbol{x}^{3^t}\right\}. \tag{19}$$

Let us remark that R and Q could be seen as a kind of generalized third-order and fourth-order moments, respectively (or, equivalently, in terms of the spherical harmonic coefficients, to the bispectrum and the trispectrum). As an example, for the particular case of uncorrelated data (i.e., $\boldsymbol{\xi} \equiv \boldsymbol{\delta}$), it is straightforward to show that $R = k_3$ and $Q = -5/2 + (3\alpha - 2)k_2 + (5/2 - \alpha)k_4$, with $k_n = \frac{1}{N}\sum_i x_i^n$.

The N-pdf $p(\boldsymbol{x}|\epsilon)$ given by equation 12 contains all the required information, on the one hand, to estimate the non-linear coupling parameter ϵ and, on the other hand, to perform a model selection (Gaussian vs. non-Gaussian). These two aspects will be studied in next Section.

Notice that the parameter α (or, equivalently, b) cannot be estimated in this framework: it just appears as an arbitrary constant in the definition of Q (equation 17): it just controls the relevance of S in Q. If one were interested in obtaining a posterior probability of the data $\boldsymbol{x}$ given both non-linear parameters (ϵ and α), then it would be necessary to expand the local non-Gaussian model beyond 0(3). However, we would always find an expression in which the non-linear parameter controlling the highest order in the expansion, acts as an arbitrary constant. For that reason, we keep terms up to 0(3). There is a discussion within the field (e.g., Okamoto & Hu 2002; Kogo & Komatsu 2006; Babich 2005; Creminelli et al. 2007) about to what extend the cubic term in the description of the weak non-linear coupling inflationary model (for which, we recall, the local non-Gaussian model in equation 1 can be seen as an approximation at large scales) is really negligible or not with respect to the quadratic contribution. For the former scenario (negligibility of the cubic term), the particular value of α becomes an irrelevant issue since, naturally, one will have that $S \ll -2 + J$ (of course, α should always satisfy the condition given in equation 5). However, for the latter case different results could be obtained, depending on the specific value for α. In the following, we will discuss parameter estimation and model selection assuming $\alpha \equiv 0$ in equation 17. However, we will explore (analyzing WMAP data and simulations) whether the condition $S \ll -2 + J$ is naturally satisfied or not.

3 STATISTICAL ANALYSIS

In this Section we aim to address two aspects very much linked one to the other: the estimation of the parameter defining the local non-Gaussian model (ϵ, Section 3.1) and the computation of some heuristic rules to decide whether a given data set is better described by the local non-Gaussian model rather than by the standard Gaussian one (Section 3.2).

3.1 Parameter estimation

The description of the full N-pdf for the non-Gaussian model proposed in Section 2 allows one to obtain an optimal estimation of the non-linear coupling parameter ϵ.

Let us recall that an *optimal estimation* of the ϵ parameter is possible, since it would be derived from the full pdf for the non-Gaussian model, $p(\boldsymbol{x}|\epsilon)$. In other words, we could obtain an unbiased and minimum variance estimator. This is possible, precisely, for the specific selection of the local non-Gaussian model in equation 3: for other physical non-Gaussian models it is not always trivial to obtain a full description of the posterior probability of the data given the parameters (at least, under realistic observational conditions like incomplete sky coverage) and shortcuts have to be taken by defining *pseudo-optimal* estimators that are, afterwards, validated with simulations.

3.1.1 Parameter estimation from the log-likelihood

We shall define the optimal estimator for the non-linear parameter as the value, $\hat{\epsilon}$, that maximizes the probability of $\boldsymbol{x}$ given ϵ. From equation 12, it is obvious that maximizing this probability is equivalent to maximize $l(\boldsymbol{x}|\epsilon)$ in equation 15. By derivation one obtains:

$$\hat{\epsilon} = \frac{R}{2Q}. \tag{20}$$

We can also estimate the error associated to $\hat{\epsilon}$ from the Fisher matrix $F_{\hat{\epsilon}} \equiv -\frac{\mathrm{d}^2 l}{\mathrm{d}\epsilon^2} = 2NQ$:

$$\sigma_{\hat{\epsilon}} = F_{\hat{\epsilon}}^{-1/2} = (2NQ)^{-1/2}. \tag{21}$$

Notice that the error on the estimation of the parameter ϵ is constant, up to the order considered in equation 3.

3.1.2 Bayesian parameter estimation

Within the Bayesian framework we can include any *a priori* information that we might have in relation to the ϵ parameter. In particular, following Bayes' theorem, the probability of ϵ given the data $\boldsymbol{x}$ read as:

$$p(\epsilon|\boldsymbol{x}) \propto p(\boldsymbol{x}|\epsilon)p(\epsilon), \tag{22}$$

where $p(\epsilon)$ is the *prior* probability function for the parameter ϵ. Of course, for the case of the local non-Gaussian model proposed in this work, there is not a clear physical motivation to choose a particular prior.

Let us however explore, as an exercise, two simple scenarios, which could be useful for more general purposes. We consider first a *uniform prior* given by:

$$p(\epsilon) \begin{cases} \frac{1}{\epsilon_{\rm M}-\epsilon_{\rm m}} & \text{if} \quad \epsilon \in [\epsilon_{\rm m}, \epsilon_{\rm M}] \\ \\ 0 & \text{otherwise} \end{cases}, \tag{23}$$

where, obviously, the range allowed to ϵ is such that $\epsilon \ll 1$ for any $\epsilon \in [\epsilon_{\rm m}, \epsilon_{\rm M}]$. For this particular case, it is trivial to show that the Bayesian estimation for the non-linear coupling parameter ($\bar{\epsilon}$) is equivalent to the one obtained via the maximum-likelihood estimation (i.e, $\bar{\epsilon} \equiv \hat{\epsilon}$) if $\hat{\epsilon} \in [\epsilon_{\rm m}, \epsilon_{\rm M}]$.

The second case we want to address corresponds to a Gaussian prior $p(\epsilon)$, described by a most probable value ϵ_* and a dispersion σ_*:

$$p(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma_*} e^{-\frac{(\epsilon-\epsilon_*)^2}{2{\sigma_*}^2}}. \tag{24}$$

By deriving the posterior probability, it is trivial to obtain the Bayesian estimation for the non-linear coupling parameter ($\bar{\epsilon}$):

$$\bar{\epsilon} = \frac{N{\sigma_*}^2 R + \epsilon_*}{2N{\sigma_*}^2 Q + 1}. \tag{25}$$

For the particular case of $\sigma_* \to 0$, i.e. a very strong prior for ϵ, peaked around ϵ_*, one trivially obtain $\bar{\epsilon} \equiv \epsilon_*$, that is, the prior dominates Bayesian estimation, leading to a most probable value for ϵ equal to the maximum value for the prior. Also trivially one finds that, for a non-informative scenario (i.e., $\sigma_* \to \infty$), $\bar{\epsilon} \equiv \hat{\epsilon} = R/2Q$.

3.2 Model selection

In this subsection we aim to calculate under which conditions (according to different model selection criteria) a given observation $\boldsymbol{x}$ is better described by a local non-Gaussian model as the one described by equation 3 with $|\epsilon| > 0$ (hereinafter H_1) rather than by a Gaussian random field described just in term of the N-point correlation function $\boldsymbol{\xi}$ (hereinafter H_0).

Some of the model selection approaches investigated in this paper have been previously applied to different astronomical/cosmological problems For instance, (Szydlowski & Godlowski 2006; Szydlowski et al 2006; Borowiec et al. 2006) applied the Akaike and the Bayesian information criteria (AIC and BIC, recpectively) to study whether astronomical data sets favored simplest models for the accelerating universe against more complex ones. This issue was also addressed by Davis et al. (2007) by analyzing the ESSENCE supernova survey data, and applying Bayesian evidence (BE) in addition to the AIC and the BIC approaches, to study. These three model selection criteria (AIC, BIC and BE) were also applied to study the impact of non-standard physical models on the Friedmann equations (Szydlowski et al. 2008). Liddle (2004) used the AIC and the BIC techniques to study, on the one hand, whether WMAP 1-year data preferred a spatially flat cosmology versus a closed one, and, on the other hand the significance of the running spectral index detected on this WMAP data release. In a posterior work (Liddle 2007), BE was added to the AIC and the BIC approaches to study the suitability of different cosmological models to the WMAP 3-year data.

Bayesian evidence (BE) is, probably, the model selection criterion that has attracted a greater interest from cosmologists during the past years. In addition to the works mentioned above (where it was compared with other model selection criteria), it has been also applied to several problems where competing cosmological models were explored. Some of these applications are the following: Mukherjee et al. (2006) followed a BE approach to study cosmological models with different matter power spectra and dark energy evolution models; Liddle et al. (2006) studied different dark energy evolving scenarios; Bridges et al. (2006, 2007b) performed a model selection on the matter power spectrum from the BE analysis of the WMAP; Bridges et al. (2007a, 2008) followed a similar approach for analyzing the WMAP compatibility with anisotropic Bianchi VII$_h$ models; Cruz et al. (2007b, 2008) used it to decide whether the WMAP cold spot was compatible or not with predictions from non-standard models like the cosmological defects; Mukherjee & Liddle (2008) studied the Planck ability to discriminate between several re-ionization models; the BE criterion was also applied (Carvalho et al. 2008; Feroz et al. 2008b) to the problem of compact source detection on microwave data; and more recently, Feroz et al. (2008a) investigated different properties of the constrained minimal supersymmetric model (mSUGRA), using WMAP data.

3.2.1 The Akaike information criterion (AIC)

The Akaike information criterion (AIC, Akaike 1973) provides with a selection index to decide among competing hypothesis, being the model associated to the lowest index the most favored one. The Akaike index corresponding to a given model or hypotheses H_i defined by p parameters and with a maximum value for the log-likelihood of $\hat{l}$ is given by:

$$\mathrm{AIC}(H_i) = 2\left(p - \hat{l}\right). \tag{26}$$

From equation 15, one can find that $\mathrm{AIC}(H_1) = 2\left(1 - N\frac{R^2}{4Q}\right)$, whereas, trivially, $\mathrm{AIC}(H_0) = 0$. Therefore, according to the AIC, the decision rule reads as:

$$\mathrm{AIC}: \begin{cases} H_0 & \text{if} \quad \frac{R^2}{Q} \leqslant \frac{4}{N} \\ \\ H_1 & \text{if} \quad \frac{R^2}{Q} > \frac{4}{N} \end{cases} \tag{27}$$

3.2.2 Bayesian information criterium (BIC)

This asymptotic bayesian criterion introduced by Schwarz (1978) is prior independent. It is based on the BIC function, that provides a measurement of the goodness-of-fit of the model to the data, taking into account the number of parameters defining the model as well as the amount of data (N):

$$\mathrm{BIC}(H_i) = \left(-2\hat{l} + p \ln N\right), \tag{28}$$

where p is the number of parameters defining the data and $\hat{l}$ is the maximum value for the log-likelihood. As for AIC, BIC provides a ranging index for competing hypothesis, where the one with the lower BIC value is the most favored one. From equation 15, one can find that (for our specific problem) $\mathrm{BIC}(H_1) = -N\frac{R^2}{2Q} + \ln(N)$ and $\mathrm{BIC}(H_0) = 0$. Therefore, according to the BIC, the decision rule reads as:

$$\mathrm{BIC}: \begin{cases} H_0 & \text{if } \frac{R^2}{Q} \leqslant \frac{2}{N}\ln N \\ \\ H_1 & \text{if } \frac{R^2}{Q} > \frac{2}{N}\ln N \end{cases} \tag{29}$$

3.2.3 Minimum description length (MDL)

MDL (e.g. Rissanen 2001) is an inference approach mostly developed during the 80s and 90s, based on the key idea that the more *regular* a given data set is, the higher is the *compression* degree to which we can code the data, and, therefore, the more we can *learn* on the properties of the data. Among many statistical applications, MDL is used to select between competing models describing the data, selecting the one allowing for a higher compression degree (which can be seen as an alternative formulation of the Occam's Razor).

For our particular case described by equation 3, the MDL measurement of compression is given by:

$$\mathrm{MDL}(\epsilon) = -\hat{l} + \frac{1}{2}\ln\left(\frac{N}{2\pi}\right) + \ln\int_{\epsilon\in\Omega} \mathrm{d}\epsilon [\det F_{\hat{\epsilon}}]^{1/2}, \tag{30}$$

where $F_{\hat{\epsilon}}$ is the Fisher matrix of $\hat{\epsilon}$. Therefore, taking into account equations 12 and 15, one can easily compute $\mathrm{MDL}(H_1)$ and $\mathrm{MDL}(H_0)$ providing a decision rule that reads as:

$$\mathrm{MDL}: \begin{cases} H_0 & \text{if } \frac{R^2}{Q} \leqslant \frac{4}{N}\ln\left(N\Omega\sqrt{\frac{Q}{\pi}}\right) \\ \\ H_1 & \text{if } \frac{R^2}{Q} > \frac{4}{N}\ln\left(N\Omega\sqrt{\frac{Q}{\pi}}\right) \end{cases}, \tag{31}$$

where Ω is the interval where ϵ is defined.

3.2.4 Generalized likelihood ratio test (GLRT)

Generalized likelihood ratio test (GLRT) is one of the most common approaches in model selection and its particular application to solve astronomical/cosmological problems has been very extensive.

The criterion established by the GLRT to accept the alternative hypothesis H_1 against H_0 is given by $p(\boldsymbol{x}|\hat{\epsilon}, H_1) > e^{\nu} p(\boldsymbol{x}|0, H_0)$ or, equivalently by:

$$\hat{l} \equiv l(\hat{\epsilon}) = N\frac{R^2}{4Q} > \nu, \tag{32}$$

where ν is an arbitrary value indicating the *strength* in choosing H_1 instead of H_0. Therefore, according to the GLRT, the decision rule reads as:

$$\mathrm{GLRT}: \begin{cases} H_0 & \text{if } \frac{R^2}{Q} \leqslant \frac{4}{N}\nu \\ \\ H_1 & \text{if } \frac{R^2}{Q} > \frac{4}{N}\nu \end{cases} \tag{33}$$

Notice that the case $\nu \equiv 1$ provides the same decision rule as the AIC (equation 27), and that $\nu \equiv \ln\sqrt{N}$ corresponds to the BIC case (equation 29).

3.2.5 Bayesian evidence (BE)

BE is defined as the average likelihood of the model H_i in the prior $p(\epsilon)$:

$$E_{H_i}(\boldsymbol{x}) = \int \mathrm{d}\epsilon\, p(\epsilon, H_i) p(\boldsymbol{x}|\epsilon, H_i), \tag{34}$$

where $p(\boldsymbol{x}|\epsilon, H_i)$ is given by equation 12. Model selection in terms of the BE grounds on the *Bayes' factor*, B_{10}:

$$B_{10}(\boldsymbol{x}) = \frac{E_{H_1}(\boldsymbol{x})}{E_{H_0}(\boldsymbol{x})}. \tag{35}$$

BE framework provides a rule to quantify how strong the decision is. In the literature it is commonly accepted the *Jeffreys' scale* (Jeffreys 1961), that provides a recipe in terms of the logarithmic Bayes' factor. Roughly speaking, it is commonly said that the evidence for H_1 against H_0 is not significant if $0 \leqslant \ln B_{10}(\boldsymbol{x}) < 1$, mild if $1 \leqslant \ln B_{10}(\boldsymbol{x}) \leqslant 3$ and strong if $\ln B_{10}(\boldsymbol{x}) > 3$.

As it was already discussed, the alternative hypothesis H_1 representing the local non-Gaussian model (equation 3) does not offer any physical motivation for a particular prior. Even thus, we study here the two particular cases already mentioned in subsection 3.1.2: the Gaussian (equation 24) and the uniform (equation 23) priors.

For the former, it is straightforward to prove that:

$$B_{10} = \frac{\sigma_{\hat{\epsilon}}}{\sqrt{\sigma_{\hat{\epsilon}}^2 + \sigma_*^2}} e^{-\frac{\epsilon_*^2}{2\sigma_*^2} + \frac{\left(\epsilon_*\sigma_{\hat{\epsilon}}^2 + \hat{\epsilon}\sigma_*^2\right)^2}{2\sigma_*^2\sigma_{\hat{\epsilon}}^2\left(\sigma_{\hat{\epsilon}}^2 + \sigma_*^2\right)}}, \tag{36}$$

whereas for the latter, one can obtain:

$$B_{10} = \sqrt{\frac{\pi}{2}}\frac{\sigma_{\hat{\epsilon}}}{\epsilon_M - \epsilon_m} e^{\frac{\hat{\epsilon}^2}{2\sigma_{\hat{\epsilon}}^2}} \left[\mathrm{erf}\left(\frac{\epsilon_M - \hat{\epsilon}}{\sqrt{2}\sigma_{\hat{\epsilon}}}\right) + \mathrm{erf}\left(\frac{\hat{\epsilon} - \epsilon_m}{\sqrt{2}\sigma_{\hat{\epsilon}}}\right)\right], \tag{37}$$

where $\hat{\epsilon}$ is the maximum-likelihood estimation for the non-linear parameter ϵ (equation 20) and $\sigma_{\hat{\epsilon}}$ is the error on this estimation (equation 21). Finally, notice that for the particular cases of $\sigma_* \gg \sigma_{\hat{\epsilon}}$ or $\epsilon_M - \epsilon_m \gg \sigma_{\hat{\epsilon}}$ in equations 36 and 37, respectively (i.e., a broad prior), one obtains:

$$B_{10} \simeq \frac{\sigma_{\hat{\epsilon}}}{\gamma} e^{\frac{\hat{\epsilon}^2}{2\sigma_{\hat{\epsilon}}^2}}, \tag{38}$$

where $\gamma = \sigma_*$ (for the Gaussian prior) or $\gamma = (\epsilon_M - \epsilon_m)/\sqrt{2\pi}$ (for the uniform prior).

4 APPLICATION TO WMAP SIMULATIONS

In this Section we aim to explore the performance of the parameter estimators and the model selection criteria described in the previous Section. We apply them to

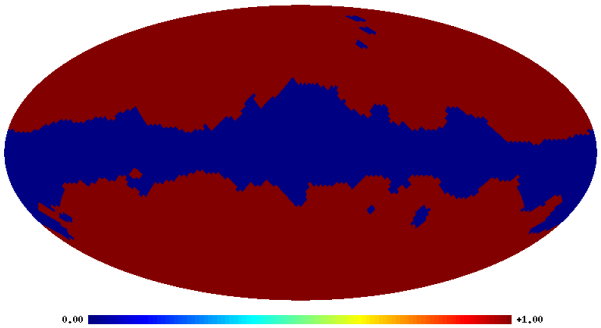

Figure 1. Mask at NSIDE=32 HEALPix resolution used in this work. It corresponds to the WMAP KQ75 mask, although the point source masking has not been considered, since the point like-emission due to extragalactic sources is negligible at the considered resolution. At this pixel resolution, the mask keeps around 69% of the sky.

CMB simulations of the WMAP 5-year data at NSIDE=32 HEALPix (Górski et al. 2005) resolution ($\approx 2^\circ$).

The procedure to generate a CMB Gaussian simulation —$(\Delta T)_G$ in equation 1— is as follows. First, using the C_ℓ obtained with the cosmological parameters provided by the best-fit to WMAP data alone (Table 6 in Hinshaw et al. 2008), we simulate WMAP observations (taking into account the corresponding beam window functions) for the Q1, Q2, V1, V2, W1, W2, W3, W4 difference assemblies at NSIDE=512 HEALPix resolution. We obtain a single co-added CMB map through a noise-weighted linear combination of the eight maps (from Q1 to W4). Weights are proportional to the inverse mean noise variance. They are independent on the position (i.e., they are uniform across the sky for a given difference assembly) and they are normalized to unity. Notice that we do not add a random noise realization to each map, since we have checked that noise plays a negligible role at the angular resolution in which we are interested in ($\approx 2^\circ$). However, we perform the linear combination of the difference assembly maps following the procedure described above, since it will be the same process that we will follow with the WMAP data.[1] Afterwards, the co-added map at NSIDE=512 is degraded down to the final resolution of NSIDE=32. Finally, a mask representing a sky coverage like the one allowed by the WMAP KQ75 mask (Gold et al. 2008) is adopted. At NSIDE=32 the mask keeps around 69% of the sky (notice that we do not consider the masking due to point sources, since at this resolution the contribution from individual extragalactic point sources is negligible, see figure 1). Let us remark that observational constraints like incomplete sky coverage can be easily taken into account by the local non-Gaussian model proposed in this work, since it is naturally defined in pixel space.

We have used 500,000 simulations of $(\Delta T)_G$, generated as described above, to estimate the correlation matrix $\boldsymbol{\xi}$ accounting for the Gaussian CMB cross-correlations. We have computed this large number of simulations to assure an accurate description of the CMB Gaussian temperature fluctuations. Additional 1,000 simulations were also generated to carry out a statistical analysis on the performance of the different parameters estimators and model selection criteria.

Each of these 1,000 $(\Delta T)_G$ simulations are transformed into $\boldsymbol{x}$ (following equations 1 and 4) to study the response of the statistical tools as a function of the non-linear ϵ parameter defining the local non-Gaussian model proposed in equation 3.

4.1 Parameter estimation

Let us first consider the estimation made via maximum-likelihood estimation (Subsection 3.1.1). As it was mentioned above, we have generated 1,000 non-Gaussian WMAP-like observations according to equation 3 for a range of values of ϵ, in particular, we have considered $\epsilon \in [0, 0.035]$, or equivalently, in terms of the most common coupling f_{NL} parameter (equation 2), we explore $f_{NL} \in [0, 500]$. We only explore positive values of the non-linear parameter, since the response of the proposed methodology does not depend on the sign of ϵ.

In figure 2 we present the $\hat{\epsilon}$ distributions obtained from the analysis of local non-Gaussian simulations. From left to right and from top to bottom, the panels show the cases: $\epsilon = 0$, 0.001, 0.002, 0.003, 0.004, 0.005, 0.010, 0.015, 0.020, 0.025, 0.030 and 0.035. Notice that, as expected, the parameter estimation is unbiased for *reasonable* values of the ϵ parameter: all the distributions are peaked around the value of ϵ used to generate the simulations. Only for $\epsilon > 0.025$ (or, equivalently, $f_{NL} \gtrapprox 350$) a bias starts to appear. This effect comes from the fact that these values of ϵ are not small enough to assure a local non-Gaussian model as the one described by equation 3, i.e., a non-Gaussian model that is a local perturbation of the underlaying CMB Gaussian signal. Also as expected (see equation 21), the width of these distributions does not depend on the particular value of ϵ, if, once more, ϵ is small enough to assure a proper expansion for the local non-Gaussian model. In this regime, we obtain, on average, $\sigma_{\hat{\epsilon}} \approx 0.004$, or, equivalently, $\sigma_{\hat{f}_{NL}} \approx 60$. Again, for $\epsilon > 0.025$ the width of the distributions starts to be slightly smaller, indicating an inadequate value of the non-linear parameter. These two effects (bias of the maximum-likelihood parameter estimation and dependence on the error on the parameter estimation) provide a natural range (at least for a pixel resolution of $\approx 2^\circ$) where the non-linear parameter is allowed to take values: $\epsilon \in [-0.025, 0.025]$.

This allowed range for ϵ can be seen as a *natural prior* $p(\epsilon)$, that could be used for performing a parameter estimation within a Bayesian framework. Obviously (as discussed in Subsection 3.1.2), the Bayesian estimation made with this uniform prior produces the same estimations for ϵ already reported from the maximum-likelihood.

Finally, we have also investigated whether the value of S in equation 17 is negligible as compared to $-2+J$, which, as discussed in Section 2, would lead to a situation where the choice of a particular value for the non-linear parameter α becomes an irrelevant problem. We found that, actually, this is not the case: on average, $|S/(-2+J)| \approx 0.7$. This implies that, first, different values of α could provide different results, not only to the ones presented in this work, but also for

[1] Co-added WMAP 5-year data is made in this way to produce a final map with a noise level smaller than, for instance, the one that could be achieved just by averaging the 8 difference assembly maps, assuring better a negligible noise contribution to the final map at resolution of $\approx 2^\circ$.

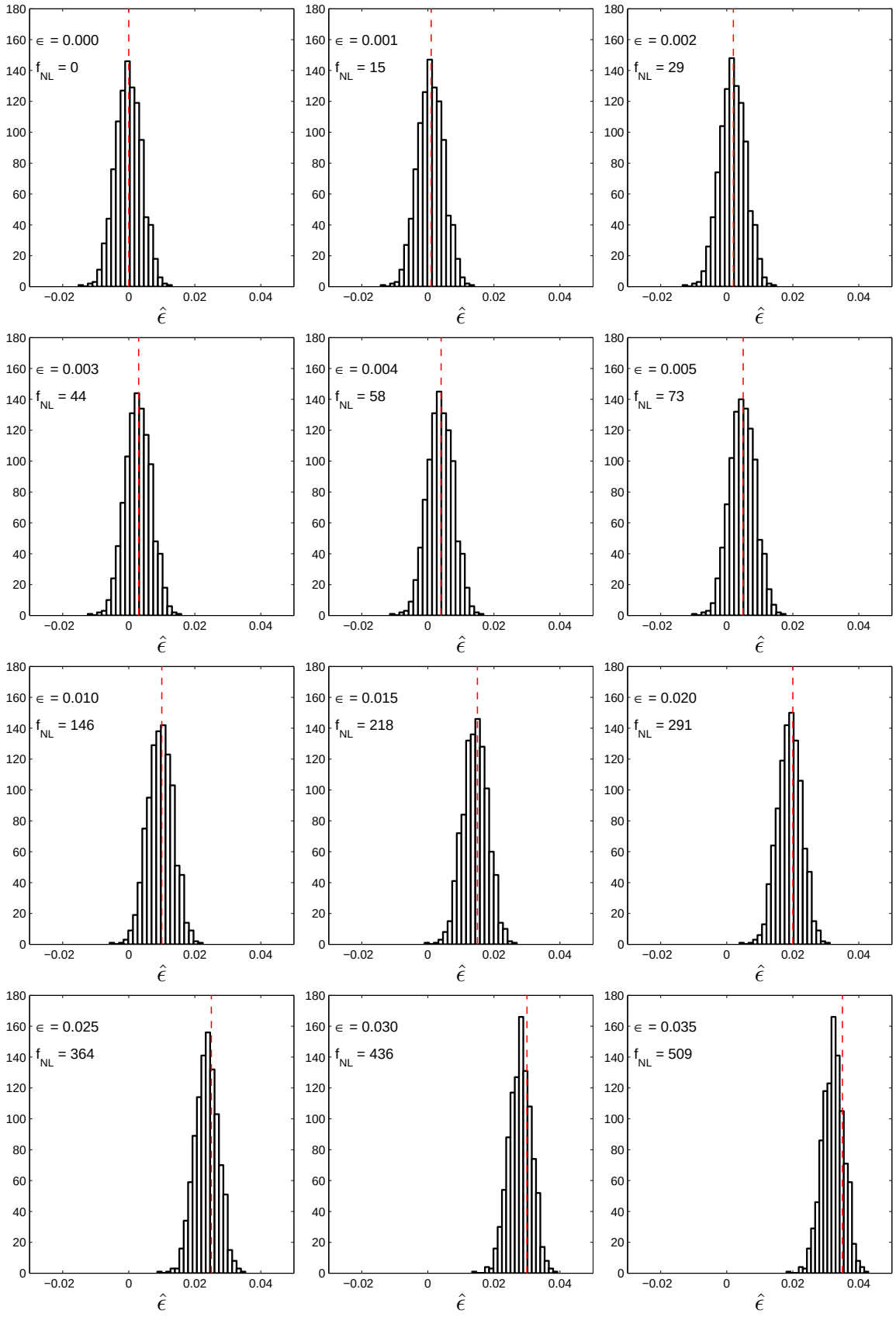

Figure 2. Distributions for the maximum-likelihood estimations of the non-linear parameter $\hat{\epsilon}$ obtained by analyzing 1,000 simulations, according to the local non-Gaussian model given in equation 3. Several values of ϵ, or, equivalently of the coupling f_{NL} parameters (both numbers are written in each panel) are explored. Vertical dashed lines indicate the value of the ϵ used to generate each set of local non-Gaussian simulations.

other works in the literature (where, we recall, it is assumed $\alpha \equiv 0$). In particular, it is trivial to show that, values of $\alpha \lesssim 0.15$ would affect the determination of the coupling parameter ϵ in $\approx 10\%$. Second, as it has been discussed by some authors (Okamoto & Hu 2002; Kogo & Komatsu 2006), it would indicate that, for the weak non-linear coupling inflationary model, the role played by cubic terms could be non negligible as compared to quadratic contributions, since, to some extent, they will contribute to the full trispectrum (as one can notice from equation 17, where α governs the role played by S). This results could be important when describing more complete non-local non-Guassian model, since it would indicate the need of including physical effects up to third order.

4.2 Model selection

We discuss which is the performance of the different model selection criteria presented in Subsection 3.2. For the particular case of BE, we assume the natural prior $p(\epsilon)$ found in the previous Subsection: $\epsilon \in [-0.025, 0.025]$. As for the case of the parameter estimation previously discussed, we have considered a set of local non-Gaussian models given by different values of the non-linear parameter.

We present the results, graphically, in figure 3. This plot consist in 6 rows and 5 columns. Each row corresponds to the results obtained for a given value of the ϵ parameter (namely, from top to bottom: $\epsilon = 0$, 0.005, 0.010, 0.015, 0.020 and 0.025). Each column refers to a model selection criterion (from left ro right: AIC, BIC, MDL, BE and GLRT). For the first column (i.e., the AIC case) we plot the distributions (obtained after analysing 1,000 simulations) of a statistical variable defined as: $\mathrm{W_{AIC}} \equiv R^2/Q - 4/N$, notice (from equation 27) that a positve value of $\mathrm{W_{AIC}}$ implies to accept H_1 against H_0. The second column accounts for the distributions of the variable $\mathrm{W_{BIC}} \equiv R^2/Q - 2\ln N/N$, which (from equation 29) also satisfies to be positive when favoring H_1. Equivalently, third column shows the distributions of $\mathrm{W_{MDL}} \equiv R^2/Q - 4/N \ln N\Omega\sqrt{Q/\pi}$, that according to equation 31 is also positive when H_1 is more likely than H_0. Fourth column provides $\ln B_{10}$, obtained from equation 37. Finally, in the fifth column we present the distributions obtained for $\nu = \left(R^2/N\right)/\left(4/N\right)$ which, according to the GLRT decision rule (equation 33) provides a measurement of the strength in accepting H_1 against H_0.

In addition to the distributions, we also plot, as an indication, a vertical line in each panel. For $\mathrm{W_{AIC}}$, $\mathrm{W_{BIC}}$ and $\mathrm{W_{MDL}}$, this vertical line separate the region where H_0 is favored (left side) from the one where H_1 is more likely (right side). The percentage of the 1,000 simulations that fall on the H_1 region is also reported at each panel. The vertical line for the $\mathrm{W_{BE}}$ statistic represents the limit for which the logarithmic Bayes' factor ($\ln B_{10}$) is greater than 1, which, in terms of the Jeffreys' rule, indicates that H_1 is, at least, mildly favored against H_0. We also provide the percentage of the simulations satisfying this condition. Finally, the vertical line for the $\mathrm{W_{GLRT}}$ statistic indicates the value of ν for which 95% of the 1,000 simulations favor H_1 instead of H_0. This value of ν is also written in the panels. Notice that, among the asymptotic model selection criteria (AIC, BIC and MDL), AIC offers a less restrictive criterion than BIC, whereas BIC behaves similarly with respect to MDL, The figure also shows that BE provides a more conservative criterion than the asymptotic methods.

5 APPLICATION TO WMAP 5-YEAR DATA

We have applied the statistical approaches described in Section 3 to WMAP 5-year data. In particular, we have analyzed a co-added CMB map generated from the global noise-weighted linear combination of the reduced foreground maps for the Q1, Q2, V1, V2, W1, W2, W3 and W4 difference assemblies (see Gold et al. 2008, for details). Weights are normalized to unity and, for each map, they are proportional to the inverse average noise variance across the sky. This operation is made at NSIDE=512 HEALPix resolution, being degraded afterwards down to NSIDE=32.

Hence, we are in the same conditions as for the analysis on simulations described in the previous Section and, therefore, the CMB cross-correlation in WMAP data is given by the $\boldsymbol{\xi}$ correlation matrix already defined in Section 4. The estimated full N-pdf of the WMAP 5-year data given the non-linear parameter ϵ is showed in figure 4 (indeed, it is given in terms of the most common $\mathrm{f_{NL}}$ parameter for allowing a better comparison with previous works). Maximum-likelihood estimation (equation 20) provides $\hat{\mathrm{f}}_{\mathrm{NL}} = 30$ with an error for the parameter (equation 21) of $\sigma_{\hat{\mathrm{f}}_{\mathrm{NL}}} = 62$ (compatible with the values obtained from simulations). Hence our WMAP 5-year data analysis reports: $\hat{\mathrm{f}}_{\mathrm{NL}} = 30 \pm 124$ at 95% CL (or, equivalently, $\hat{\epsilon} = 0.019 \pm 0.078$ at 95%). This result is compatible with similar works in the literature reporting WMAP compatibility with Gaussian hypothesis. However, let us remark that this estimation is more efficient than previous ones at similar angular resolution, since it provides a smaller error bar. For instance, Curto et al. (2007) performed a Gaussianity test on the WMAP data, at the same HEALPix resolution, although in a smaller region of the sky (16% instead of the 69% considered in this work). Their $\mathrm{f_{NL}}$ estimator was based on the three Minkowski functionals, providing an error bar of ≈ 200. The expected $\sigma_{\hat{\mathrm{f}}_{\mathrm{NL}}}$ provided by our maximum-likelihood estimator for a similar observed region would be $\sigma_{\hat{\mathrm{f}}_{\mathrm{NL}}} \approx 124$, i.e., $\approx 40\%$ smaller than the one obtained by the Minkowski functionals. This result was expected since, as it was already discussed, the maximum-likelihood estimation of the non-linear parameter is *optimal*, given the local non-Gaussian model in equation 1.

As it has been mentioned above, this result shows WMAP data compatibility with the Gaussian hypothesis, since $\mathrm{f_{NL}} \equiv 0$ can not be rejected at any significant confidence level. Of course, same conclusions are obtained from the model selection criteria described in Subsection 3.2. Neither AIC, BIC, MDL nor BE criteria select H_1 against H_0, whereas GLRT would favor H_1 under the very weak condition for ν in equation 33 of $\nu \approx 0.1$ (which implies a likelihood ratio of ≈ 1.1)

Finally, let us remark that, as it also happened for the simulations analyzed in Section 4, the contribution of the S term to the value of Q (see equation 17) is not negligible, in particular, $|S/(-2+J)| = 0.74$. This indicates that, as it was already mentioned, the cubic term in the model given by equation 3 is not *naturally* negligible as compared to the quadratic term and, therefore, α should be chosen small enough (like the case $\alpha \equiv 0$ considered in this work).

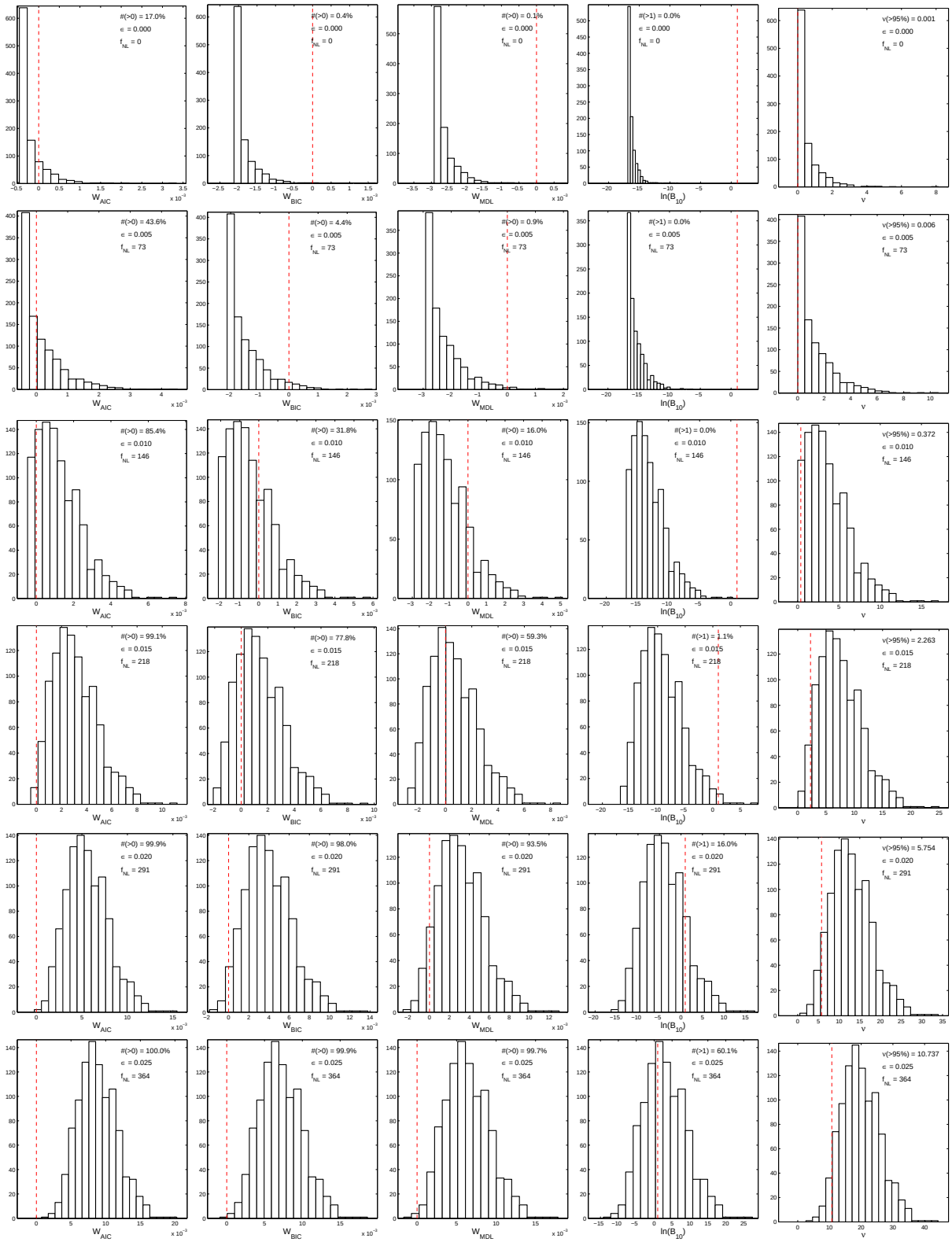

Figure 3. From left ro right, columns show the distribution for several statistics referred to different model selection criteria: AIC, BIC, MDL, BE and GLRT. From top to bottom, results for local non-Gaussian models for different non-linear parameters are give: $\epsilon = 0$, 0.005, 0.010, 0.015, 0.020 and 0.025. Panels in columns 1, 2, 3 and 4 present a vertical line, separating the region where H_0 and H_1 are preferred (to the left and to the right of the vertical line, respectively). The vertical line on the last column represents the value for the ν parameter for which 95% of the non-Gaussian simulations are more likely described by H_1 rather than H_0.

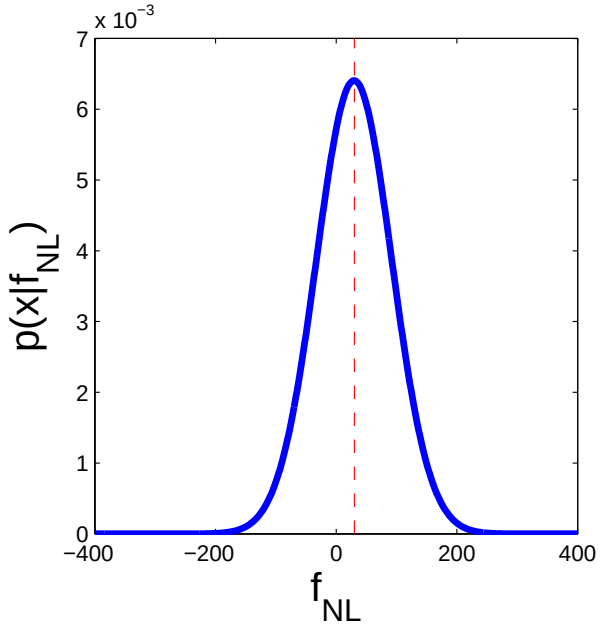

Figure 4. This curve represent the probability given in equation 12, i.e. the full pdf of the WMAP 5-year data $\boldsymbol{x}$ given the non-linear parameter ϵ (or, equivalently, f_{NL}). Vertical dotted line marks the maximum-likelihood estimation $\hat{f}_{NL} = 30$.

6 CONCLUSIONS

We have presented a parametric non-Gaussian model for the CMB temperature fluctuations. The non-Gaussian model is a local perturbation (up to third order) of the standard CMB Gaussian field which recovers (for the case of $b \equiv 0$ in equation 1) an approximative form of the weak non-linear coupling inflationary model (e.g. Komatsu et al. 2001; Liguori et al. 2003) at scales larger than the horizon scale at the recombination time (i.e. above the degree scale). For this model, we are able to build the posterior probability of the data given the non-linear parameter ϵ (see equation 3), from which, in principle, an *optimal estimator* (i.e., unbiased and with minimum variance) can be derived. Analytical expressions for the maximum-likelihood estimation of the non-linear parameter ($\hat{\epsilon}$) and its associated error ($\sigma_{\hat{\epsilon}}$) are derived. In addition, we also discuss an alternative Bayesian estimation (in terms of the posterior probability of the non-linear parameter), for the hypothetical case in which we might have some prior information for ϵ. As an example, two cases are addressed: a non-informative (i.e, uniform) and a Gaussian priors. We also investigate an issue very much linked to the parameter estimation: the model selection. Indeed, we discuss several well known techniques to perform hypotheses test, like the Akaike information criterion (AIC, Akaike 1973), the Bayesian information criterion (BIC, Schwarz 1978), the minimum description length (MDL, Rissanen 2001), the generalized likelihood ratio test (GLRT) and the Bayesian evidence (BE). We derive analytical expressions, for the particular local non-Gaussian model proposed in this work, for all these model selection techniques.

The performance of both, parameter estimators and model selection criteria, are investigated by analyzing non-Gaussian simulations, as they could be observed by WMAP. We check that the maximum-likelihood estimation provides an unbiased and efficient estimation of the non-linear parameter defining the deviations from Gaussianity. We find that, for the HEALPix resolution considered in this work (NSIDE=32), results are consistent up to a value of $\epsilon = 0.025$, which approximately corresponds (at the Sachs-Wolfe regime) to a value of the the non-linear coupling parameter $f_{NL} \approx 350$. This parameter is the one commonly used to described the weak non-linear coupling inflationary model (e..g Komatsu et al. 2001). We also find that, among the model selection criteria, AIC is the asymptotic method that provides the less restrictive decision rule, whereas, on the other hand, MDL is the most strict one. We also find that BE, for a uniform prior given by $\epsilon \in [-0.025, 0.025]$, is even more restrictive than MDL.

The proposed methodology is applied to WMAP 5-year data. We obtain a value for the non-linear coupling parameter of $\hat{f}_{NL} = 30 \pm 124$ at 95% CL. This result provides a more efficient estimation than previous works in the literature, at the same angular scales. For instance, comparing with the work by Curto et al. (2007) using Minkowski functionals, we can infer that the maximum-likelihood error bar is $\approx 40\%$ smaller than the one obtained with those geometrical estimators. Application of model selection criteria to WMAP data confirms that standard hypothesis of Gaussianity is favored against the alternative hypothesis of non-Gaussianity, for the specific local model proposed in this work, and for the adopted resolution of $\approx 2^\circ$.

Finally, we would like to comment that, currently, we are extending the technique based on the N-pdf presented in this work, to deal with a more realistic non-local non-Gaussian model, where higher resolution CMB data are considered, including as well the effect of anisotropic noise.

ACKNOWLEDGEMENTS

The authors thank Andrés Curto, R. Belén Barreiro and Enrique Martínez-González for useful comments and discussion. We acknowledge partial financial support from the Spanish Ministerio de Ciencia e Innovación project AYA2007-68058-C03-02. PV also acknowledges financial support from the Ramón y Cajal programme. PV thanks to the CNR Istituto de Scienza e Tecnologie dell'Informazione (ISTI, Pisa) for their warm hospitality during his research stays in March and June 2008. JLS acknowledge partial financial support by the Spanish MEC and thanks the CNR ISTI in Pisa for their hospitality during his sabbatical leave. The authors acknowledge the computer resources, technical expertise and assistance provided by the Spanish Supercomputing Network (RES) node at Universidad de Cantabria. We acknowledge the use of Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for it is provided by the NASA Office of Space Science. The HEALPix package was used throughout the data analysis Górski et al. (2005).

REFERENCES

Abramo L.R., Bernui A., Ferreira I.S., Villela T., Wuensche C.A., 2006, Phys. Rev. D, 74, 063506

Ackerman L., Carroll S.M., Wise M.B. 2007, Phys. Rev. D, 75, 083502

Akaike H., 1973, Proceed. of the 2nd International Symposium on Information Theory (eds. Pertov B.N. & Czaki F), Akad. Kiado, Budapest, 267

Babich D., 2005, Phys. Rev. D., 72, 043003

Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, Phys. Rep., 402, 103

Bernui A., Villela T., Wuensche C.A., Leonardi R., Ferreira I., 2006, A&A, 454, 409

Bernui A., Mota B., Rebouças M.J., Tavakol R., 2007, A&A, 464, 479

Bielewicz P., Eriksen H.K., Banday A.J., Górski K.M., Lilje P.B., 2005, ApJ, 635, 750

Böhmer C.G., Mota D.F., 2008, Phys. Lett. B, 663, 168

Borowiec A., Godlowski W., Szydlowski M., 2006, Phys. Rev. D, 74, 043502

Bridges M., Lasenby A.N., Hobson M.P., 2006, MNRAS, 369, 1123

Bridges M., McEwen J.D., Lasenby A.N., Hobson M.P., 2007a, MNRAS, 377, 1473

Bridges M., Lasenby A.N., Hobson M.P., 2007b, MNRAS, 381, 68

Bridges M., McEwen J.D., Cruz M., Lasenby A.N., Hobson M.P., Vielva P., Martínez-González E., 2008, MNRAS, 390, 1372

Cabella P., Liguori M., Hansen F.K., Marinucci D., Matarrese S., Moscardini L., Vittorio N., 2005, MNRAS, 358, 684

Carvalho P., Rocha G., Hobson M.P., 2008, MNRAS, submited (preprint arXiv0802.3916)

Cayón L., Martínez-González E., Argüeso F., Banday A.J., Górski K.M., 2003, MNRAS, 339, 1189

Cayón L., Jin J., Treaster A., 2005, MNRAS, 362, 826

Chiang L.-Y., Naselsky P.D., Verkhodanov O. V., 2003, ApJ, 590, 65

Chiang L.-Y., Naselsky P. D., 2006, Int. J. Mod. Phys. D, 15, 1283

Coles P., Dineen P., Earl J., Wright D., 2004, MNRAS, 350, 989

Copi C.J., Huterer D., Starkman G.D., 2004, Phys. Rev. D, 70, 043515

Creminelli P., Nicolis A., Senatore L., Tegmark M., Zaldarriaga M., 2006, Journal of Cosmology and Astro-Particle Physics, 5, 4

Creminelli P., Senatore L., Zaldarriaga M., 2007, Journal of Cosmology and Astro-Particle Physics, 3, 19

Cruz M., Martínez-González E., Vielva P., Cayón L., 2005, MNRAS, 356, 29

Cruz M., Tucci M., Martínez-González E., Vielva P., 2006, MNRAS, 369, 57

Cruz M., Cayón L., Martínez-González E., Vielva P., Jin J., 2007a, ApJ, 655, 11

Cruz M., Turok N., Vielva P., Martínez-González E., Hobson M.P., 2007b, Science, 318, 1612

Cruz M., Martínez-González E., Vielva P. Diego J.M., Hobson M.P., Turok N., 2008, MNRAS, 390, 913

Curto A., Aumont J., Macías-Pérez, Martínez-González E., Barreiro R.B., Santos D., Désert F.-X., Tristram M., 2007, A&A, 474, 23

Curto A., Martínez-González E., Mukherjee P., Barreiro R.B., Hansen F.K., Liguori M., Matarrese S., 2008, MNRAS, in press

Davis T.M. et al., 2007, ApJ, 666, 716

de Oliveira-Costa A., Tegmark M., Zaldarriaga M., Hamilton A., 2004, Phys. Rev. D, 69, 063516

Donoghue E.P., Donoghue J.F., 2005, Phys. Rev. D, 71, 043002

Eriksen H.K., Hansen F.K., Banday A.J., Górski K.M., Lilje P.B., 2004a, ApJ, 605, 14

Eriksen H.K., Novikov D.I., Lilje P.B., Banday A.J., Górski K.M., 2004b, ApJ, 612, 64

Eriksen H.K., Banday A.J., Górski K.M., Lilje P.B., 2005, ApJ, 622, 58

Eriksen H.K., Banday A.J., Górski K.M., Hansen F.K., Lilje P.B., 2007, ApJ, 660, L81

Feroz F., Allanach B.C., Hobson M., Abdus Salam S.S., Trotta R., Weber A.M., 2008a, Journ. of High Energy Phys., 10, 64

Feroz F., Hobson M.P., Zwart J.T.L., Sounders R.D.E., Grainge K.J.B., 2008b, MNRAS, submited (preprint arXiv0811.1199)

Freeman P.E., Genovese C.R., Miller C.J., Nichol R.C., Wasserman L., 2006, ApJ, 638, 1

Gordon C., 2007, ApJ, 656, 636

Gold B., 2008, ApJS, in press

Górski K.M., Hivon E., Banday A.J., Wandelt B.D., Hansen F.K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Gott J.R., Colley W.N., Park C.-G., Park C., Mugnolo C., 2007, MNRAS, 377, 1668

Groeneboom N.E., Eriksen H.K., 2008, ApJ, in press

Hansen F.K., Cabella P., Marinucci D., Vittorio N., 2004a, ApJ, 607, L67

Hansen F.K., Banday A. J., Górski K.M., 2004b, MNRAS, 354, 641 0, 063004

Hikage C., Matsubara T., Coles P., Liguori M., Hansen F. K., Matarrese S., 2008, MNRAS, 389, 1439

Himmetoglu B., Contaldi C.R., Peloso M., 2008a, (preprint arXiv0809.2779)

Himmetoglu B., Contaldi C.R., Peloso M., 2008a, (preprint arXiv0812.1231)

Hinshaw G. et al., 2008, ApJS, in press

Jaffe T.R., Banday A.J., Eriksen H.K., Górski K.M., Hansen F.K., 2006a, ApJ, 629, L1

Jaffe T.R., Hervik S., Banday A.J., Górski K.M., 2006b, ApJ, 644, 701

Jaffe T.R., Banday A.J., Eriksen H.K., Górski K.M., Hansen F.K., 2006c, preprint (arXiv:astro-ph/0606046v1)

Jeffreys H., 1961, Theory of Probability (3rd edition) Oxford university Press

Katz G., Weeks J., 2004, Phys. Rev. D, 70, 063527

Kogo N., Komatsu E., 2006, Phys. Rev. D., 73, 083007

Komatsu E., Spergel D.N., 2001, Phys. Rev. D, 63, 063002

Komatsu E., et al., 2003, ApJS, 148, 119

Komatsu E., et al., 2008, ApJS, in press

Land K., Magueijo J., 2005a, MNRAS, 357, 994

Land K., Magueijo J., 2005b, Phys. Rev. Lett., 95, 071301

Land K., Magueijo J., 2007, MNRAS, 378, 153

Liguori M., Matarrese S., Moscardini L., 2003, ApJ, 597,

57
Liddle A., Lyth D. H., 2000, *Cosmological inflation and large-scale structure*, Cambridge University Press
Liddle A., 2004, MNRAS, 351, 49
Liddle A., 2007, MNRAS, 377, 74
Liddle A., Mukherjee P., Parkinson D., Wang Y., 2006, Phys. Rev. D, 74, 123506
Martínez-González E., Cruz M., Cayón L., Vielva P., 2006, New Astron. Rev., 50, 875
McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2005, MNRAS, 259, 1583
McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2006, MNRAS, 371, 50
Monteserín C., Barreiro R.B., Vielva P., Martínez-González, Hobson, M.P., Lasenby A.N., 2008, MNRAS, 387, 209
Mukherjee P., Wang Y., 2004, ApJ, 613, 51
Mukherjee P., Parkinson D., Liddle A., 2006, ApJL, 638, 51
Mukherjee P., Liddle A., 2008, MNRAS, 389, 231
Okamoto T., Hu W., 2002, Phys. Rev. D., 66, 063008
Park C.-G., 2004, MNRAS, 349, 313
Pietrobon D., Amblard A., Balbi A., Cabella P., Cooray A., Mirinucci D., 2008, Phys. Rev. D, 10, 3504
Räth C., Schuecker P., Banday A.J., 2007, MNRAS, 380, 466
Reissanen J., 2001, IEEE Transactions on Information Theory, 47, 5
Schwarz G., 1978, Ann. Stat., 6, 461
Schwarz D.J., Starkman G.D., Huterer D., Copi C.J., 2004, Phys. Rev. Lett., 93, 221301
Spergel D.N. et al., 2007, ApJS, 170, 377
Szydlowski M., Godlowski W., 2006, Phys. Lett. B, 633, 427
Szydlowski M., Kurek A., Krawiec A., 2006, Phys. Lett. B, 642, 171
Szydlowski M., Godlowski W., Stachowiak T., 2008, Phys. Rev. D, 77, 043530
Tojeiro R., Castro P.G., Heavens A.F., Gupta S., 2006, MNRAS, 365, 265
Vielva P., Martínez-González E., Barreiro R.B., Sanz J.L., Cayón L., 2004, ApJ, 609, 22
Vielva P., Wiaux Y., Martínez-González E., Vandergheynst P., 2006, New Astron. Rev., 50, 880
Vielva P., Wiaux Y., Martínez-González E., Vandergheynst P., 2007, MNRAS, 381, 932
Wiaux Y., Vielva P., Martínez-González E., Vandergheynst P., 2006, Phys. Rev. Lett., 96, 151303
Wiaux Y., Vielva P., Barreiro R.B., Martínez-González, Vandergheynst P., 2008, MNRAS, 385, 939
Yadav A.P.S., Wandelt B.D., 2008, Phys. Rev. Lett., 100, 181301